Effect of functional-group distribution on the structure of a polymer in non-aqueous solvent

ARTICLE TEMPLATE

Effect of functional-group distribution on the structure of a polymer in non-aqueous solvent

Rui F. G. Apóstolo[a], Philip J. Camp[a], Beatrice N. Cattoz[b], Peter J. Dowding[b], and Andrew D. Schwarz[b]

[a]School of Chemistry, University of Edinburgh, David Brewster Road, Edinburgh EH9 3FJ, Scotland; [b]Infineum UK Ltd, P.O. Box 1, Milton Hill, Abingdon OX13 6BB, U.K.

ARTICLE HISTORY
Compiled 3 July 2018

ABSTRACT
The sizes and structures of isolated functionalised polymers in a hydrocarbon solvent are studied using atomistic molecular dynamics simulations and Monte Carlo simulations of coarse-grained chains. A specific functionalised polyethylene-polypropylene random copolymer in n-heptane is studied using atomistic simulations. The functional groups contain aromatic and polar groups, and 8 of them are distributed on an 8 kDa polymer backbone in several different ways. It is shown that the radius of gyration and the end-to-end distance depend sensitively on the functional-group distribution. A random distribution of functional groups gives the most compact polymer structure, but other distributions gives values up to 50% larger; the largest values are when the functional groups are split evenly between both ends of the polymer. This is shown to be due to the association of the polar, and hence solvophobic, functional groups. A coarse-grained bead-spring model is then studied that includes solvophilic beads (representing unfunctionalised units) and solvophobic beads (representing functionalised units). Monte Carlo simulations are used to survey functional-group concentration and distribution. The results show that the collapse of a polymer with increasing solvophobicity depends sensitively on the distribution of different beads. Form factors are presented for both the atomistic and coarse-grained models, and are analysed as if they were experimental scattering measurements. The apparent radii of gyration are in good agreement with those determined directly from simulation.

KEYWORDS
Functionalised polymers; molecular dynamics; Monte Carlo; coarse graining

1. Introduction

Recent advances in synthetic strategies mean that functionalised polymers (FPs) can be produced with a variety of compositions, structures, and architectures [1–3]. The diversity of chemical functionalities that can be incorporated with backbone polymer structures allows many applications of FPs in chemical synthesis, biological science, chemical engineering, medicine, agriculture, food science, optical science, sensor technology, and the energy sector [4, 5]. In many cases, the functionality of the polymer relies directly on very specific chemical interactions and reactions of the functional groups, such as in sensing applications. In other cases, the properties of the FP de-

CONTACT P. J. Camp. Email: philip.camp@ed.ac.uk

pend more on the physical structure, which is of course dictated by the functional groups. For example, in lubricants, polymer functionalisation can be used to tune the properties of viscosity modifiers, friction modifiers, and dispersants (of soot in oil) [6, 7].

The connections between the molecular architecture of an FP, and the physical properties of the FP (or a solution of the FPs), are complex. For FPs in solution, the interactions between the polymer backbone groups, the functional groups, and the solvent molecules can be resolved into effective (strictly many-body) forces between the backbone and functional groups, with the solvent degrees of freedom having been integrated out. For instance, if an FP possesses a solvophilic backbone and solvophobic functional groups, then one can anticipate repulsive interactions between backbone groups because they are in good-solvent conditions, and attractive effective interactions between the functional groups because they are in bad-solvent conditions.

The link between polymer architecture and bulk-phase structure is well established in the cases of block copolymers and graft polymers [8]. Many different arrangements of two or more polymer blocks can be realised in the laboratory. The amount of 'blockiness' controls the structures, symmetries, and domain sizes in thermally or chemically quenched polymer melts, and this can often be thought of as a kind of microphase separation of the blocks. The phase behaviour of such materials can be rationalised using self-consistent field theory [9–12] or particle simulations of coarse-grained models [13, 14].

Many functionalisation routes give random distributions of functional groups on the backbone, but there are techniques for producing other, more regular distributions [15–18]. For instance, the functional groups may be evenly spaced on the backbone, or concentrated in the middle or at one end, or split between both ends. If the FPs are only sparsely functionalised, though, then their behaviour can be quite different from that of block copolymers, because the functional groups do not have to be concentrated in large domains. Of course, taking this to an extreme, the structures, properties, and functions of biological polymers such as proteins are dictated by the numbers and precise locations of residues on a polypeptide backbone of, on average, 150 amino acids [19]. Biological polymers obviously present a much more complicated and subtle problem than FPs, given the complexity and specificity of the interactions between residues (electrostatic interactions, hydrogen bonding, hydrophobic interactions, disulfide bridges, etc.) and the variety of secondary structures. Nonetheless, even for this complicated problem, modelling proteins as amphiphilic polymers can uncover some important features, for example, how the distribution of hydrophilic and hydrophobic residues controls the aggregation of proteins [20].

This work is focused on the size of a functionalised polyethylene/polypropylene (PE/PP) random copolymer in a non-aqueous solvent (n-heptane); the PE/PP backbone is in good-solvent conditions. (Future experimental work will involve small-angle neutron scattering (SANS) studies, and deuterated heptane is a convenient means of improving contrast between solute and solvent.) The functional groups contain aromatic and polar chemical species, and hence in an aliphatic solvent, there are attractive interactions between them; the functional groups are proprietary (belonging to Infineum UK Ltd), and so more details cannot be given here. The main task is to determine how the effective interactions and the distribution of functional groups on the backbone affect the physical properties of a FP. The simplest property of a single polymer is its size, measured either by the radius of gyration (R_{g}) or its end-to-end distance (R_{ee}). Of course, chemical details are important, and the precise way in which backbone and functional groups interact cannot easily be predicted. But once some-

thing is known about the effective interactions between these groups, then a much simpler coarse-grained model can be sought that captures the essential features of the effective interactions. A lot of computational effort is expended in determining effective, coarse-grained interactions in polymer molecules [21, 22], and in favourable cases, the interaction potentials are transferable [23]. But even the most primitive models, such as chains of soft-spheres, can yield some valuable insights, and it is then possible to survey a wider range of molecular variables at a fraction of the computational cost.

The physical dimensions of a single FP in solution are studied here using large-scale atomistic molecular dynamics (MD) simulations, and Monte Carlo simulations of a primitive bead-spring model. The rest of the article is organised as follows. In Section 2, the atomistic MD simulations (2.1) and MC simulations of coarse-grained chains (2.2) are described. The results are presented in Section 3 and Section 4 concludes the paper.

2. Models and methods

2.1. Atomistic molecular dynamics simulations

All molecular parameters are chosen to be representative of proprietary material belonging to Infineum UK Ltd. The model FP was represented with the OPLS-AA force field, which gives a consistent set of parameters for both united-atom and explicit-atom representations [24–26]; non-polar aliphatic and aromatic carbon-hydrogen units (CH, CH_2, CH_3) were represented as united atoms, and all other atoms, including polar hydrogens, were represented explicitly.

The FP backbone was constructed from a sequence of 151 PE groups (4236 Da) and 89 PP groups (3745 Da) giving a 53%:47% PE:PP ratio by mass. The total backbone mass was 8 kDa. Functional groups were attached to 8 tertiary propylene carbons, as it is known that functionalisation occurs preferentially at these sites [27, 28]. There were 21 united and explicit atoms in each functional group. Sequences of unfunctionalised PE units and PP units, and functionalised PP units, were generated according to five different functional-group distributions: spread out evenly along the backbone ('even'); randomly distributed along the backbone ('random'); all in the middle of the backbone ('middle'); all at one end of the backbone ('one end'); and evenly split between both ends of the backbone ('both ends'). Due to the computational demand, only one random configuration was considered, with the distribution of the 81 unfunctionalised PP groups (U) and 8 functional PP groups (F) being $U_{14}FU_{10}FU_6FU_{11}F_2U_4FU_{10}FU_{25}FU$. Hence, the number of unfunctionalised PP units separating functional groups ranged from 0 to 25. In the even distribution, there were either 11 or 12 unfunctionalised PP units between functional groups, and this represents a significant difference which manifests itself in the measured molecular properties.

The FP was first simulated in vacuum with NVT MD at $T = 298$ K in a cubic box of side 100 Å for about 0.1 ns, in order to relax its structure. The FP was then solvated with 4286 n-heptane molecules using Packmol [29, 30], and equilibrated under the same NVT conditions for 2.5 ns. Then the simulation was switched to NPT conditions at $P = 1$ atm and $T = 298$ K, and runs of 400 ns were carried out. In general, the structural parameters for each FP conformation showed no systematic drift after 100 ns, and statistics were gathered in the remaining 300 ns. All MD simulations were carried out using LAMMPS [31, 32] with periodic boundary conditions applied, the Nosé-Hoover thermostat/barostat, and the velocity Verlet algorithm with timesteps

of either 0.5 fs (relaxation and solvation/equilibration in NVT ensemble) or 2.0 fs (equilibration/production in NPT ensemble).

The radius of gyration (R_g) was calculated as an RMS average using the formula

$$R_\mathrm{g}^2 = \frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle r_{ij}^2 \rangle \tag{1}$$

including all N united and explicit FP atoms, where r_{ij} is the separation between atoms i and j without the minimum-image convention applied [33, 34]. This was carried out by reconstructing the polymer in empty space, working along the backbone monomer-by-monomer. In experiments, R_g can be determined from the low-wavevector behaviour of the form factor $P(q)$ obtained from SANS or small-angle X-ray (SAXS) scattering. To simulate this procedure, the isotropic form factor was computed using the formula

$$\frac{P(q)}{P(0)} = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\langle \frac{\sin(qr_{ij})}{qr_{ij}} \right\rangle \tag{2}$$

defined so that $P(q)/P(0) = 1$, and again the interatomic distances were calculated without periodic boundary conditions applied. Here, each united and explicit atom is assumed to have the same scattering cross section.

2.2. Coarse-grained Monte Carlo simulations

A bead-spring model [35–38] was constructed with two types of bead: type-1 beads represent unfunctionalised units in good-solvent conditions; and type-2 beads represent functionalised units in bad-solvent conditions. All bead-bead interactions were expressed in terms of the Lennard-Jones (LJ) potential

$$u_\mathrm{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] \tag{3}$$

where ϵ and σ are the energy and range parameters, respectively. All interactions involving type-1 beads were given by the Weeks-Chandler-Andersen potential

$$u_{11}(r) = u_{12}(r) = \begin{cases} u_\mathrm{LJ}(r) - u_\mathrm{LJ}(r_0) & r \leq r_0 \\ 0 & r > r_0 \end{cases} \tag{4}$$

where $r_0 = 2^{1/6}\sigma$ is the position of the minimum in the LJ potential, and $u_\mathrm{LJ}(r_0) = -\epsilon$. Interactions between type-2 beads [$u_{22}(r)$] were given by a similar potential to that in Equation (4) but with r_0 replaced by a cut-off distance $r_\mathrm{c} = 2.5\sigma$. Beads were bonded with their neighbours in the polymer chain by a simple harmonic potential $u_\mathrm{bond}(r) = \frac{1}{2}K_\mathrm{bond}r^2$ with $K_\mathrm{bond} = 10\epsilon\sigma^{-2}$. In combination with the non-bonded potentials given above, this results in a ground-state separation between a pair of beads of 1.10σ, slightly less than r_0. The harmonic potential was used instead of the familiar finitely extensible nonlinear elastic potential [35] because its comparative softness speeded up equilibration in simulations. Figure 1(a) shows the non-bonded, bonded, and total interactions between a pair of beads with either repulsive interactions (u_{11},

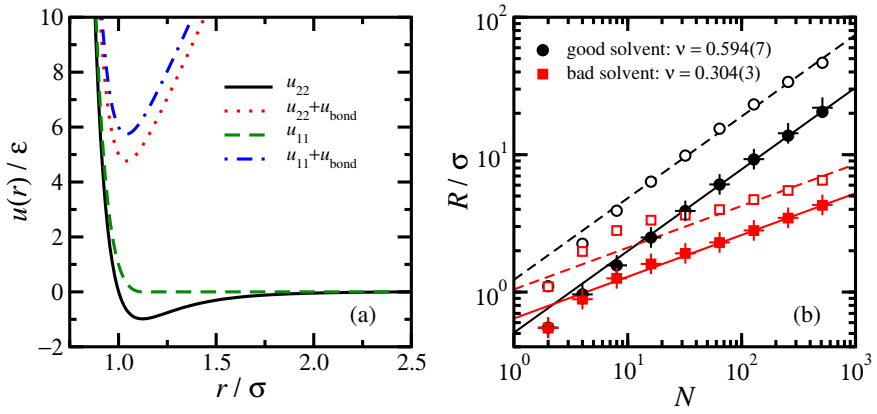

Figure 1. (a) Coarse-grained interaction potentials involving beads in good solvent ($u_{11} = u_{12}$) and between beads in bad solvent (u_{22}). The total interactions between bonded beads are $u_{ij} + u_{\rm bond}$, where $u_{\rm bond}$ is a harmonic potential (not plotted separately). (b) Radius of gyration ($R_{\rm g}$, filled symbols) and end-to-end distance ($R_{\rm ee}$, unfilled symbols) from MC simulations of homopolymers with N type-1 beads (good solvent, circles) or N type-2 beads (bad solvent, squares). The crosses show the values of $R_{\rm g}$ determined by fitting the low-q portion of the form factor $P(q)$. Also shown are fits to the scaling law $R_{\rm g} \sim N^{\nu}$ with $N \geq 32$ (type 1) and $N \geq 128$ (type 2), and corresponding fits to $R_{\rm ee}$ with the same values of ν.

u_{12}) or attractive interactions (u_{22}). The composition of a polymer chain is defined by the type-2 fraction x_2. Single chains with various compositions were studied using conventional NVT Metropolis MC simulations with an average acceptance rate of 50% for single-particle moves, and around 10^8 attempted moves per particle for both the equilibration and production runs [33, 34]. The dimensionless temperature was set to $T^* = k_{\rm B}T/\epsilon = 1$, where $k_{\rm B}$ is Boltzmann's constant. For the random configuration, results for 8 independent realisations were averaged.

3. Results

3.1. Atomistic MD simulations

Figure 2 shows final configurations from atomistic MD simulations of the unfunctionalised polymer, and FPs with five different distributions. Note that there are 21 united and explicit atoms in each functional group. In all cases, the polymers are in coil-like conformations, and from simulation movies, there is a high degree of molecular flexibility. Hence, there are large fluctuations in the apparent sizes of polymers, and there is no single snapshot that characterises each polymer.

Figure 3 shows results for $R_{\rm g}$ and $R_{\rm ee}$ of the FPs. Figure 3(a) shows the evolution of the instantaneous value of $R_{\rm g}$ over 400 ns, and that there is no systematic drift after about 100 ns. Note that the results are shifted up by multiples of 30 Å for clarity. Figure 3(b) shows probability distributions from the final 300 ns. Note that in each case, the distribution is broad, meaning that the width is comparable to the average. This shows that the polymers are very flexible, and that they explore a wide range of conformations on the timescale of 100s of nanoseconds. The results depend sensitively on the functional-group distribution. With an even distribution, there is no strongly preferred value in the range 20 Å $\leq R_{\rm g} \leq$ 50 Å. With a random distribution, the most probable value is around 25 Å, and with all of the functional groups in the middle of the polymer, the most probable value is about 30 Å. When all of the functional groups are at one end of the polymer, there is a broad distribution of $R_{\rm g}$ values. When the

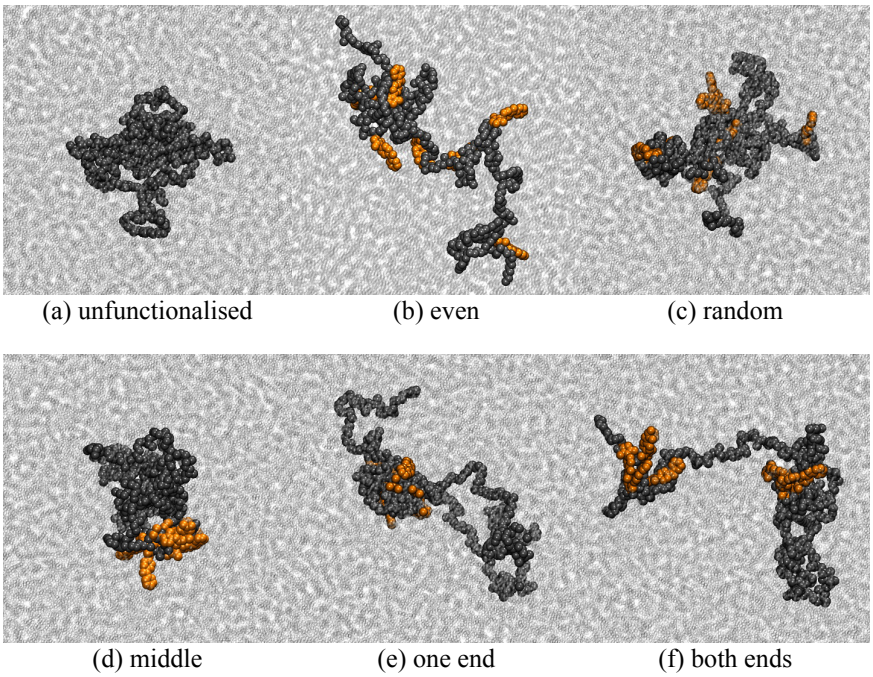

Figure 2. Final configurations of FPs (shown in space-filling representation) in n-heptane (shown in stick representation) from atomistic MD simulations. The CH_2 and CH_3 united atoms in the PE/PP backbone are shown in black, and the functional-group united and explicit atoms are shown in orange.

functional groups are split between both ends of the polymer, the most probable value of $R_g \simeq 40$ Å is far greater than with any of the other distributions. Slightly different behaviour is seen in Figures 3(c) and (d) showing the corresponding results for R_{ee}, in that there are less-pronounced differences between the even, random, middle, and one-end distributions. Note that in Figure 3(c), the results are shifted up by multiples of 60 Å for clarity. There is no systematic drift apparent in the data, but for the both-ends distribution, there is a step-like decrease at about 275 ns. It would take much longer simulations or biasing methods to assess the statistical weights of these conformational manifolds, so it will simply be noted that there are two (or more) of them.

The root-mean-square averages of R_g and R_{ee} for the FPs over the time interval 100 ns $\leq t \leq$ 400 ns are shown in Table 1. Results for the unfunctionalised PE/PP backbone are also shown. The statistical uncertainties were calculated using a blocking procedure [34]. The averages for the FPs are also shown as horizontal lines over the corresponding time intervals in Figure 3(a) and (c). In the case of the both-ends distribution, average values of R_{ee} in the time intervals 100 ns $\leq t \leq$ 275 ns and 275 ns $\leq t \leq$ 400 ns are also given in Table 1 and Figure 3(c). The essential point is that, in terms of both R_g and R_{ee}, the ordering of the FP sizes is

$$\text{random} < \text{middle} < \text{even} < \text{one end} < \text{both ends}.$$

Most of the differences in R_g are at least 2 Å, except for the even and one-end distributions which differ by about 1 Å. The unfunctionalised polymer cannot be fitted into this sequence because it has a smaller molecular mass than the FPs.

To get insight on the trend for the FPs, Figure 4 shows probability distribution

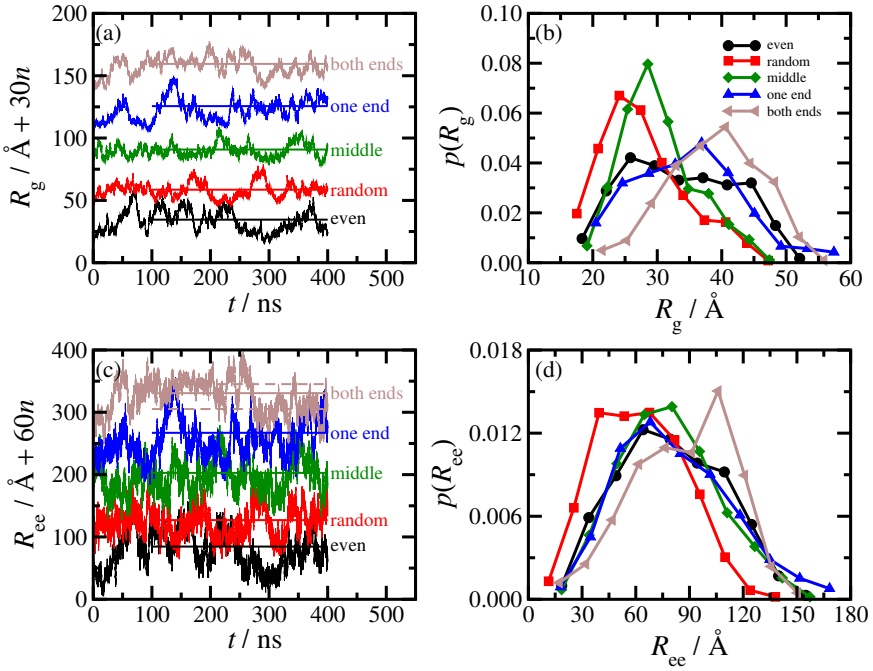

Figure 3. (a) Instantaneous values of R_g at time t from atomistic MD simulations; the results are shifted up by multiples of 30 Å for clarity. (b) Probability distributions $p(R_g)$ over the time interval 100 ns $\leq t \leq$ 300 ns. (c) Instantaneous values of R_{ee} at time t; the results are shifted up by multiples of 60 Å for clarity. (d) Probability distributions $p(R_{ee})$ over the time interval 100 ns $\leq t \leq$ 300 ns. In (a) and (c), the solid lines indicate averages over the interval 100 ns $\leq t \leq$ 400 ns. In (c), the dashed lines indicate averages over the intervals 100 ns $\leq t \leq$ 275 ns (upper) and 275 ns $\leq t \leq$ 400 ns (lower).

Table 1. Radius of gyration (R_g) and end-to-end distance (R_{ee}) from atomistic MD simulations. R_g was determined by direction calculation [Equation (1)] and from analysis of the form factor [Equation (5)].

Distribution	R_g/Å Direct	R_g/Å From $P(q)$	R_{ee}/Å
Unfunctionalised	29.6 ± 1.7	30.5 ± 0.1	65.7 ± 5.3
Even	34.5 ± 2.7	34.5 ± 0.1	84.6 ± 8.7
Random	28.6 ± 2.1	29.6 ± 0.1	67.1 ± 5.3
Middle	30.8 ± 1.3	31.6 ± 0.1	82.8 ± 5.9
One end	35.6 ± 2.2	34.9 ± 0.1	87.1 ± 7.1
Both ends	39.5 ± 1.7	40.1 ± 0.1	90.7 ± 8.9[a]

[a] Averages over different time intervals are as follows: 100 ns $\leq t \leq$ 275 ns, $R_{ee} = (105.2 \pm 2.6)$ Å; 275 ns $\leq t \leq$ 400 ns, $R_{ee} = (65.2 \pm 4.8)$ Å. See Figure 3(c).

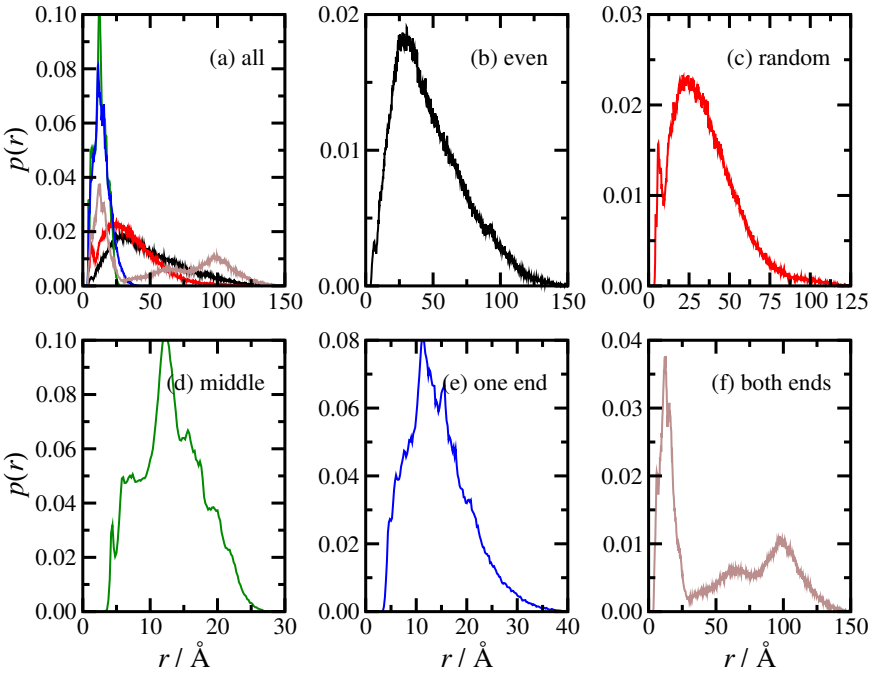

Figure 4. Probability distribution functions $p(r)$ for the centres of mass of the functional groups with different distributions from atomistic MD simulations: (a) all results; (b) even; (c) random; (d) middle; (e) one end; (f) both ends.

functions $p(r)$ for the centres of mass of the functional groups, normalised so that $\int_0^\infty p(r)\mathrm{d}r = 1$. The variations in R_g and R_ee are subtle, but the spatial correlations between the functional groups show marked differences. The correlations between functional groups are broadly similar for the even and random distributions, broadly similar for the middle and one-end distributions, and very different for the both-ends distribution. For the even distribution [Figure 4(b)] there is a peak at about 25 Å and a very long tail up to 150 Å. For the random distribution [Figure 4(c)] the correlations are more pronounced in the range up to about 75 Å, and in addition, there is a short-distance peak corresponding to functional groups in contact with one another. For the middle and one-end distributions [Figure 4(d) and (e), respectively] the functional groups are clearly clustered in a domain of size 25–30 Å. Finally, the both-ends distribution shows two clear features: short-range correlations between functional groups at one end; and long-range correlations between functional groups at opposite ends of the chain, arising from the connecting polymer backbone.

These results shed some light on the measured order of the radius of gyration and end-to-end distance. Firstly, the random distribution gives the shortest polymer size, in terms of both R_g and R_ee. This is due to two main factors: firstly, there are strong, short-range contacts between functional groups nearby on the polymer backbone, giving the short-distance peak in $p(r)$, and keeping the polymer pinned in a compact coil-like conformation with small R_g; there is clearly a knock-on effect on R_ee, which is significantly smaller than with the other distributions. The even distribution has larger R_g and R_ee because the uniform distance between functional groups is too large to allow clustering; this also explains the absence of a short-distance peak in $p(r)$. The middle and one-end distributions also result in increased polymer sizes, as compared to the random distribution. This is due to there being some association between functional groups in the interior of the coil, but there is

nothing to connect opposite ends of the polymer chain. That the middle distribution gives a smaller size than the one-end distribution suggests that it is like two small unfunctionalised chains (in good solvent) attached to the ends of a functionalised segment (in bad solvent), resembling a dumbbell. Both R_g and R_{ee} show that the overall structure is more compact than with the one-end distribution. Finally, the both-ends distribution gives by far the largest values of R_g and R_{ee}. While the 4 functional groups at each end can form compact structures, Figure 3(d) shows that there is a very low probability of association between functional groups at opposite ends of the chain. Moreover, the average value of R_{ee} after the crossover at $t \simeq 275$ ns is close to that of the random distribution. Although the functional groups are solvophobic, the strength of attraction is insufficient to pin the rest of the polymer chain into a ring-like conformation. The plots of $p(R_g)$ and $p(R_{ee})$ in Figure 3(b) and (d) show that there is a finite, but low, probability of the polymers having very compact conformations, and that the polymers sample very wide ranges of R_g and R_{ee}. Hence, the polymers have sufficient opportunity during the 400 ns simulations to explore the most probable states.

To complete this section, the form factors of the polymers have been computed [Equation (2)] so that there is a connection with future SANS/SAXS studies. The results are shown in Figure 5 on conventional log-log plots. All results for the FPs are compared in Figure 5(a). Apparently, there are only small differences between the different functional-group distributions, but a detailed analysis shows that the trends reported above can be detected from $P(q)$. The low-q behaviour could be analysed using the Guinier approximation $P(q) \approx 1 - (qR_g)^2/3$, but it turns out that the form factor for a simple model of a linear polymer gives a better fit over a wider range of q. There are no exact analytical results for polymers with excluded-volume interactions, [39] but the following function for a linear Gaussian chain is adequate in the current case [40–42], and of course it obeys the Guinier approximation at low q:

$$\frac{P(q)}{P(0)} = \frac{2}{(qR_g)^4}\left[e^{-(qR_g)^2} - 1 + (qR_g)^2\right]. \qquad (5)$$

The fits were made in the range $q \leq 0.1$ Å^{-1} and are shown over the whole q range in Figure 5(b)–(f). At low wavevectors, the fits are sufficiently good to give rather precise values for the R_g. The fitted results are given in Table 1, and they are in excellent agreement with the values obtained by direct calculation, with deviations of 1 Å or less.

3.2. Coarse-grained MC simulations

MC simulations of chains with $N = 2$–512 type-1 beads (good solvent) or type-2 beads (bad solvent) were carried out. Figure 1(b) shows the radius of gyration (R_g) and end-to-end distance (R_{ee}). Results for R_g are shown from both direct calculation [Equation (1)] and by fitting to $P(q)$ (see below). This plot shows that, with large values of N, the non-trivial scaling relation $R \sim N^\nu$ is obeyed. In good-solvent conditions, the fit to R_g ($N \geq 32$) gives $\nu = 0.594 \pm 0.007$, which is in good agreement with the accepted Flory exponent $\nu = 0.5880 \pm 0.0015$ [43, 44]. In bad-solvent conditions, the fit to R_g ($N \geq 128$) gives $\nu = 0.304 \pm 0.003$ which is in moderate agreement with the trivial value $\nu = \frac{1}{3}$ expected for a compact globule. The end-to-end distance is statistically more noisy due to it depending only on the positions of two beads, but the results are

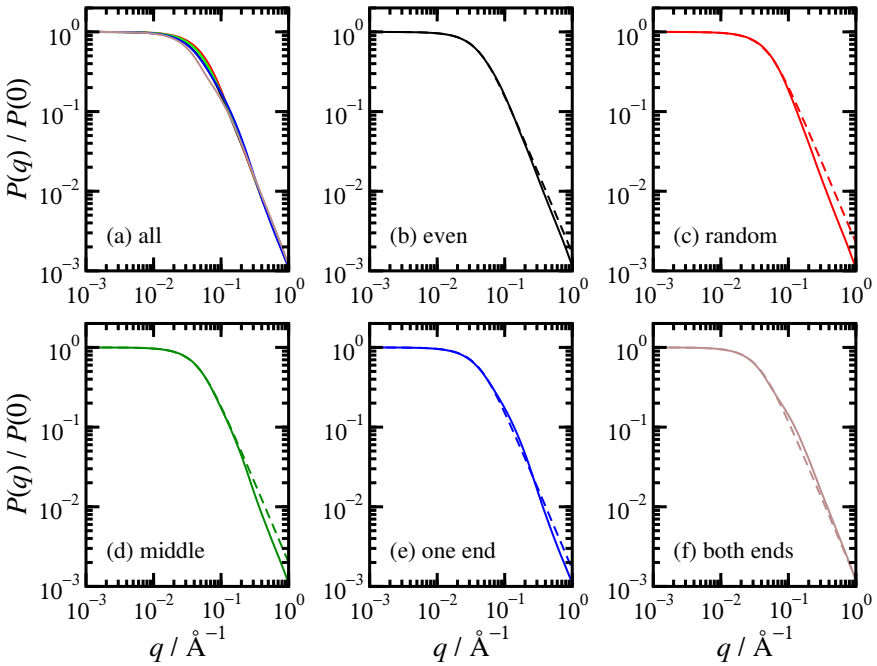

Figure 5. Form factors from atomistic MD simulations with different functional-group distributions: (a) all results; (b) even; (c) random; (d) middle; (e) one end; (f) both ends. The solid lines are output from the simulations, and the dashed lines are fits using Equation (5) ($q \leq 0.1$ Å^{-1}).

still consistent with the scaling laws; the fits in Figure 1(b) were constrained by the values of ν obtained from R_g.

Chains containing $N = 128$ beads are in the scaling regime, and so chains of this length were studied containing both type-1 and type-2 beads. The type-2 fractions studied here are $x_2 = 0$, $\frac{1}{16}$, $\frac{1}{8}$, $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$, $\frac{7}{8}$, $\frac{15}{16}$, and 1, so that there are at least 8 type-2 beads in a heteropolymer. With these compositions, there is an integer number of domains in the even distribution, and with $x_2 \leq \frac{1}{2}$ ($x_2 \geq \frac{1}{2}$), single type-2 (type-1) beads are spaced equally along the chain. Final snapshots of chains with $x_2 = \frac{1}{4}$ and even, random, middle, one-end, and both-ends distributions are shown in Figure 6. These images show that the type-2 beads cluster by analogy with the association of functional groups observed in the atomistic MD simulations. With the middle, one-end, and both-ends distributions, the type-2 domains are tightly clustered. With the even and random distributions, only small numbers of type-2 beads are in mutual contact. The both-ends distribution switches between conformations where the type-2 beads at opposite ends are clustered and unclustered. Two representative snapshots are shown with $R_{ee} \simeq 3\sigma$ and 19σ, which are close to the peaks in the corresponding probability distribution, to be described next.

As with the atomistic simulations, the natural fluctuations in polymer conformation are very significant. This is shown in the probability distributions $p(R_g)$ and $p(R_{ee})$ presented in Figure 7. Two cases are presented: $x_2 = \frac{1}{4}$ and $x_2 = \frac{1}{2}$. The results for $x_2 = \frac{1}{4}$ show that for all of the functional-group distributions except the both-ends case, $p(R_g)$ and $p(R_{ee})$ are unimodal. In the both-ends case, however, the probability distributions are bimodal. $p(R_{ee})$ is expected to be Gaussian, while $p(R_g)$ is more complicated [45], but for the purposes of illustration, both are fitted with a sum of two Gaussians. The fraction of the low-R_g/low-R_{ee} peak is about 0.3. The results

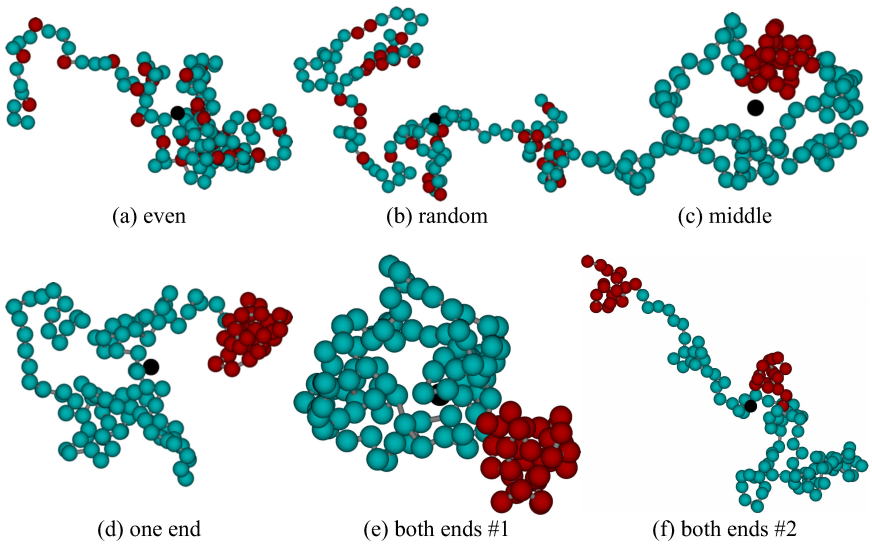

Figure 6. Snapshots from MC simulations of coarse-grained chains with $x_2 = \frac{1}{4}$ and different distributions: (a) even; (b) random; (c) middle; (d) one end; (e) both ends with $R_{ee} \simeq 3\sigma$; (f) both ends with $R_{ee} \simeq 19\sigma$. Type-1 beads are shown in blue and type-2 beads are shown in red. The centre of mass is denoted by the black dot.

for $x_2 = \frac{1}{2}$ show that in the both-ends case, the increase in type-2 beads causes the polymer to collapse. Comparing Figure 7(a) and (b) also shows that R_g for the middle distribution also changes significantly on increasing the type-2 bead fraction. The variations of the averages with x_2 are discussed next.

Figure 8 shows R_g and R_{ee} as functions of x_2. Runs for each distribution included chains with $x_2 = 0$ and 1, and these values were averaged to give the values plotted in Figure 8. Two sets of values for R_g are plotted in Figure 8(a), one by direct calculation [Equation (1)], and one by fitting to $P(q)$ (see below); the results are, of course consistent. In general, both R_g and R_{ee} decrease with increasing x_2 because the polymer collapses as it moves from good-solvent conditions to bad-solvent conditions. Focusing on R_g first, the data are all very similar near $x_2 = 0$ and $x_2 = 1$, where the effects of the minority beads are small. The largest differences between the results are near $x_2 = \frac{1}{2}$, where R_g increases in the order

$$\text{both ends} < \text{middle} < \text{one end} < \text{random} < \text{even}$$

which differs from that seen in the atomistic MD simulations. The both-ends distribution is vastly different in the two cases: in the atomistic model, the inherent rigidity of the polymer backbone precludes association between the two ends, meaning that R_g is large; in the coarse-grained model, the chain has the flexibility to fold up. Some tests were done on models with a bending potential of the Kratky-Porod form [46] $u_{\text{bend}}(\theta) = K_{\text{bend}}(1 - \cos\theta)$, where $K_{\text{bend}} \leq 5\epsilon$ and θ is the angle between successive bond vectors in the chain, but all that happened was that the polymer collapse in all models shifted to higher values of x_2, and without changing the order of the R_g values. Turning to R_{ee}, there are strong differences in behaviour between the different distributions. For the even and random distributions, R_{ee} shows a very similar behaviour to that of R_g. R_{ee} for the middle and one-end distributions are similar to one another, but vary with x_2 in a different way to R_g. Finally, R_{ee} for the both-ends distribution shows that there is a sharp binding between the two type-2 domains at

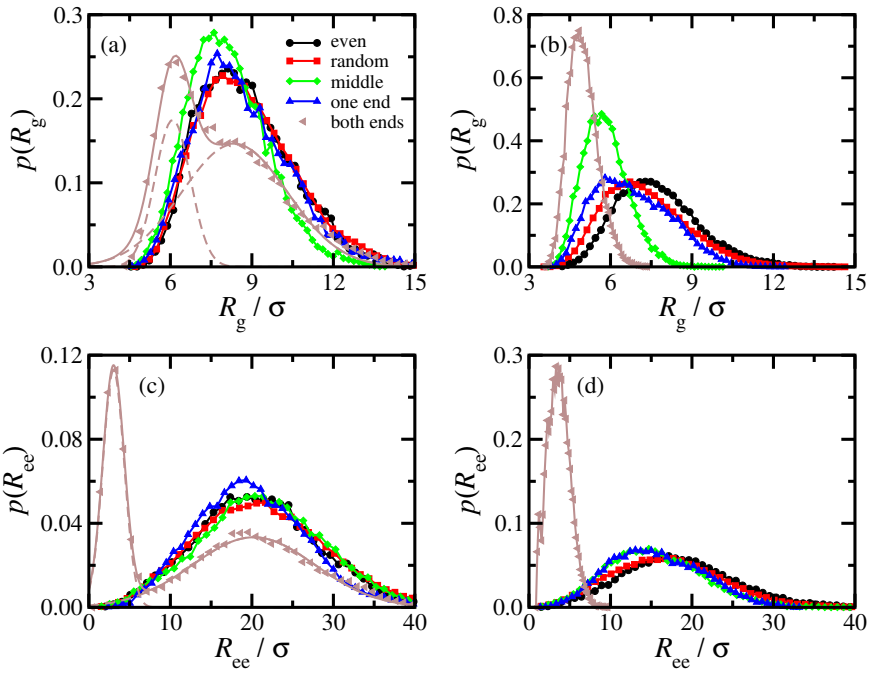

Figure 7. Probability distributions of R_g [(a) and (b)] and R_{ee} [(c) and (d)] from MC simulations of coarse-grained chains with $x_2 = 1/4$ [(a) and (c)] and $x_2 = 1/2$ [(b) and (d)]. In most cases, the results are shown as points connected with lines. The exceptions are the results for the both-ends distribution with $x_2 = 1/4$ [(a) and (c)]: the MC results are shown as points, and the lines are fits to a sum of two Gaussians; the individual contributions are shown with dashed lines and the sums are shown with solid lines.

$x_2 \simeq \frac{1}{4}$. So, overall, the values of R_{ee} are not necessarily slaved to the values of R_g, and the orderings are quite different. The discord between the atomistic modelling and the coarse-grained modelling is hardly surprising, as the bead-spring model is a very naïve one. Nonetheless, some interesting trends can be discerned from the results.

R_g and R_{ee} for each distribution were fitted with the function

$$R(x_2) = \frac{1}{2}(A+B) + \frac{1}{2}(B-A)\,f(x_2) \quad (6)$$

$$f(x_2) = C\tanh\left[D(x_2-y_2) + E(x_2-y_2)^2\right] \quad (7)$$

where $A = R(0)$, $B = R(1)$, and $f(x_2)$ is an interpolating function such that $f(y_2) = 0$. The values of C and E are constrained by the fitted values of D and y_2 so that $f(0) = -1$ and $f(1) = 1$. Hence, the fitted function interpolates between A and B, y_2 is the crossover point between expanded and collapsed configurations, and D measures the steepness of the curve at y_2. The fitted values of D and y_2 are shown in Table 2. Of particular note are the values of y_2, which are measures of the compositions at which the chains crossover from good-solvent to bad-solvent behaviour. Looking first at R_g, the values of y_2 increase in the order

$$\text{both ends} < \text{middle} < \text{random} < \text{one end} < \text{even}.$$

which is almost the same order as R_g at except that the random and one-end distributions are switched. This is reflected in the different slopes of the curves near y_2, given by the parameter D; in fact, the random distribution has the largest value of

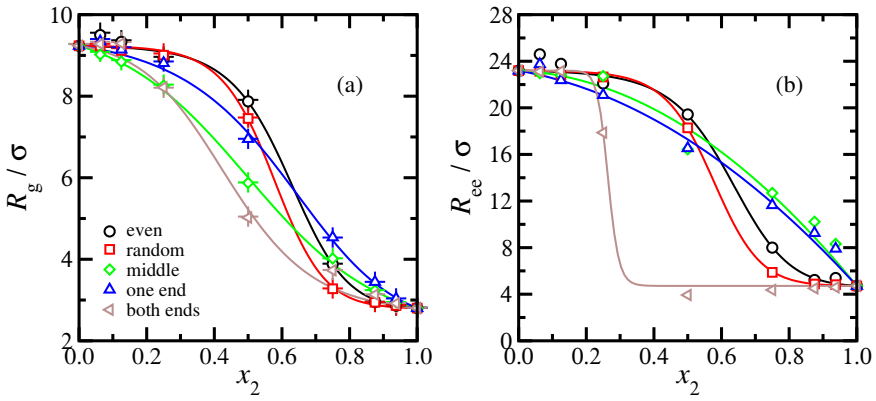

Figure 8. (a) Radius of gyration (R_g) and (b) end-to-end distance (R_{ee}) as functions of type-2 bead fraction x_2 from MC simulations of coarse-grained chains of $N = 128$ beads. The points are from simulations and the curves are fits using Equation (6). In (a), the crosses show the values of R_g determined by fitting the low-q portion of the form factor $P(q)$.

Table 2. Parameters D and y_2 from fitting Equation (6) to R_g and R_{ee} of coarse-grained chains of $N = 128$ beads. A and B are fixed at the limiting values for homopolymers. The other parameters $C = [\tanh(Dy_2 - Ey_2^2)]^{-1}$ and $E = D(2y_2 - 1)/(1 - 2y_2 + 2y_2^2)$ are fixed so that $R(0) = A$ and $R(1) = B$. Where given, the figures in brackets are the fitting uncertainties in the final digit.

	R_g/Å				R_{ee}/Å			
Distribution	A	B	D	y_2	A	B	D	y_2
Even	9.24	2.80	5.7(4)	0.619(9)	23.2	4.71	5.6(7)	0.62(2)
Random	9.24	2.80	6.49(3)	0.57(1)	23.2	4.71	7.0(3)	0.573(4)
Middle	9.24	2.80	2.6(2)	0.499(7)	23.2	4.71	0.0061	0.70(2)
One end	9.24	2.80	3.0(2)	0.606(9)	23.2	4.71	0.0045	0.66(2)
Both ends	9.24	2.80	3.9(6)	0.44(2)	23.2	4.71	31	0.26

D, followed closely by the even distribution. So the picture is that the random and even distributions 'snap' into collapsed states on increasing x_2, whereas the middle, one-end, and both-ends distributions show a more gradual crossover in R_g. Turning to R_{ee}, the even and random distributions show quite similar values of y_2 and D. The curvatures of the functions near y_2 are very different for the middle and one-end distributions, the values of D are very small and with meaninglessly large fitting errors, and so the values are only given to 2 significant figures in Table 2. Finally, the both-ends distribution is anomalous because of the association between opposite ends of the chain; the fitted function is not really able to describe the simulation results, and again the fit parameters are given to 2 significant figures. R_{ee} increases slightly in the range $x_2 \geq \frac{1}{2}$ because as the proportion of type-2 beads increases, it is less energetically necessary for the ends of the chain in a collapsed conformation to be in contact, while entropy favours the ends being dissociated from one another.

To complete this section, some examples of form factors for coarse-grained chains are shown in Figure 9. Results are shown for the middle distribution, which shows the most linear variation of R_g with increasing x_2 [Figure 8(a)]. Figure 9(a) shows $P(q)$ for chains with $x_2 = 0$, $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$, and 1. With increasing x_2, the shoulder in $P(q)$ shifts to higher wavevector, reflecting the decreasing value of R_g, and there is an increase in structure in the range $1 \leq q\sigma \leq 2\pi$ due to the aggregation of type-2 beads into liquid-like clusters with nearest neighbour distances close to the value of σ. Figure 9(b)–(f) shows the component parts $P_{11}(q)$ and $P_{22}(q)$, calculated separately by restricting the

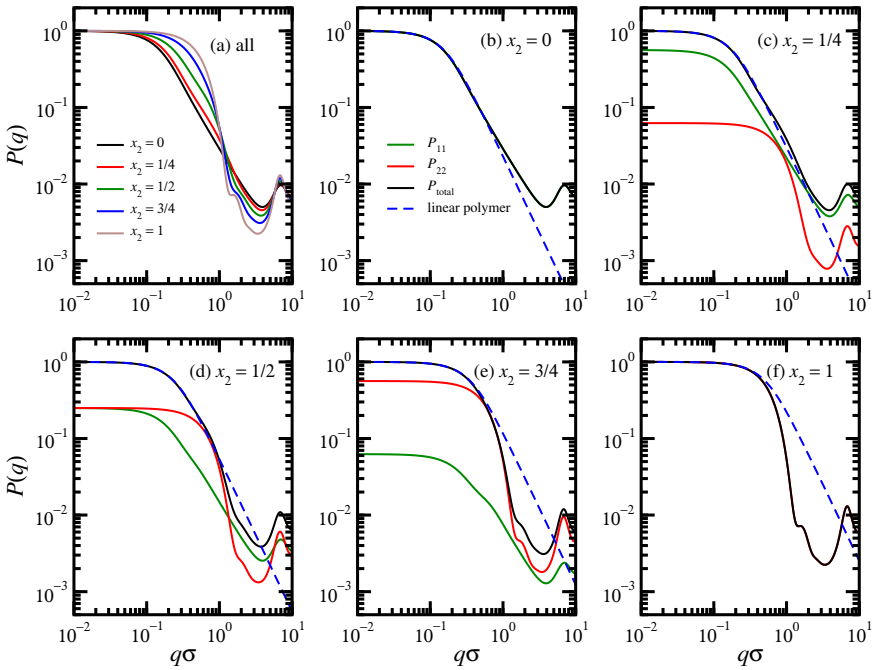

Figure 9. Form factors from MC simulations of chains with $N = 128$ beads and the middle distribution of type-2 beads: (a) all results; (b) $x_2 = 0$; (c) $x_2 = \frac{1}{4}$; (d) $x_2 = \frac{1}{2}$; (e) $x_2 = \frac{3}{4}$; (f) $x_2 = 1$.

sums in Equation (2) to type-1 and type-2 beads, respectively; $P(q)$ is then given by the sum of these, plus a cross term P_{12} (not shown). This decomposition of $P(q)$ shows that the type-1 beads retain expanded, good-solvent structures with a form factor similar to that of the Gaussian chain. The type-2 beads form dense clusters with a quite distinct form factor. This behaviour could be discerned experimentally by isotopic enrichment (deuteration) of one or other of the two types of monomer unit. Finally, Figures 9(b)–(f) also show the form factors of the Gaussian chain with R_g fitted in the range $q\sigma \leq 0.1$. The R_g values are also shown in Figure 8 and they are, of course, consistent with those from direct calculations.

4. Conclusions

In this work, the effects of functional-group distribution on the size of a FP in solution have been studied using atomistic MD simulations and MC simulations of a coarse-grained model. The atomistic simulations were of a solvophilic polymer functionalised with solvophobic groups, dissolved in *n*-heptane. It was found that both the radius of gyration and the end-to-end distance of the FP can vary by up to 50%, depending on the functional-group distribution. A random distribution gives the most compact conformation, while a both-ends distribution gives the most expanded conformation. This was shown to be due to the association of the functional groups. The random configuration has a range of separations of consecutive functional groups along the backbone, and this allows strong association between adjacent groups, and weaker association between groups widely separated on the chain. With the both-ends distribution, functional groups at opposite ends of the chain do not associate, presumably because of the entropic cost of pinning the two ends together, and restricting the

number of conformations available to the connecting backbone. The observed dependence of polymer size on functional-group distribution is clearly dictated by specific interactions between functional groups that are near to each other, and the inherent bending and torsional rigidity of the polymer backbone.

The importance of chemical detail was highlighted by the results of a very simple, bead-spring model of a FP, consisting of solvophilic beads (representing unfunctionalised units) and solvophobic beads (representing functionalised units). The dependence of polymer size on the distribution of solvophobic beads was quite different from that observed in atomistic simulations. Nonetheless, using such a simple model, it was easy to survey the properties of FPs with increasing solvophobicity. It was found that the crossover from expanded to collapsed conformations depends sensitively on the functional-group distribution. Moreover, unlike in the chemically detailed simulations, the end-to-end distance is not slaved to the radius of gyration, mainly due to the flexibility of the coarse-grained model, and probably also the missing steric bulk of the functionalised units. Whereas the solvophobic beads in the coarse-grained models simply cluster into liquid-like drops, the association of the functional groups in the real FP is strongly controlled by the rigidity and bulk of the backbone. In fact, this justifies the investment of substantial amounts of computing time to study atomistic models of FPs in explicit solvent!

The next step is to examine the effects of functional-group distribution on the surface-adsorption properties of FPs, the aggregation of several FPs, and the response of these to the application of shear flow. This work is in progress.

Acknowledgements

The authors are privileged to contribute this work to the special issue of Molecular Physics in honour of Daan Frenkel's 70^{th} birthday.

Funding

This work was supported by Infineum UK Ltd through a studentship to R.F.G.A. We gratefully acknowledge the support of NVIDIA Corporation through the donation of a GeForce Titan Xp GPU used in this research.

References

[1] N.K. Boaen and M.A. Hillmyer, Chem. Soc. Rev. **34**, 267–275 (2005).

[2] N.M.G. Franssen, J.N.H. Reek and B. de Bruin, Chem. Soc. Rev. **42**, 5809–5832 (2013).

[3] E. Blasco, M.B. Sims, A.S. Goldmann, B.S. Sumerlin and C. Barner-Kowollik, Macromolecules **50**, 5215–5252 (2017).

[4] A. Akelah and A. Moet, *Functionalized Polymers and their Applications* (Springer, Netherlands, 1990).

[5] A. Akelah, *Functionalized Polymeric Materials in Agriculture and the Food Industry* (Springer, New York, 2013).

[6] M.D. Gieselman, Mannich post-treatment of PIBSA dispersants for improved dispersion of EGR soot 2012, US Patent 8,324,139.

[7] M.R. Sutton, W.R.S. Barton, D. Price and M.C. Davies, Dispersant Viscosity Modifiers 2012, US Patent App. 13/382,351.

[8] P.G. Khalatur and A.R. Khokhlov, Soft Matter **9**, 10943–10954 (2013).
[9] M.W. Matsen, Macromolecules **28**, 5765–5773 (1995).
[10] M.W. Matsen and F.S. Bates, J. Chem. Phys. **106**, 2436–2448 (1997).
[11] F. Drolet and G.H. Fredrickson, Phys. Rev. Lett. **83**, 4317–4320 (1999).
[12] G.H. Fredrickson, V. Ganesan and F. Drolet, Macromolecules **35**, 16–39 (2002).
[13] P. Košovan, J. Kuldová, Z. Limpouchová, K. Procházka, E.B. Zhulina and O.V. Borisov, Macromolecules **42**, 6748–6760 (2009).
[14] J. Kuldová, P. Košovan, Z. Limpouchová, K. Procházka and O.V. Borisov, Collect. Czech. Chem. Commun. **75**, 493–505 (2010).
[15] D.A. Shipp, J. Macromol. Sci., Polym. Rev. **45**, 171–194 (2005).
[16] S. Pfeifer and J.F. Lutz, J. Am. Chem. Soc. **129**, 9542–9543 (2007).
[17] K. Matyjaszewski and N.V. Tsarevsky, Nature Chem. **1**, 276–288 (2009).
[18] M. Benaglia, M. Chen, Y.K. Chong, G. Moad, E. Rizzardo and S.H. Thang, Macromolecules **42**, 9384–9386 (2009).
[19] C.B. Anfinsen, Science **181**, 223–230 (1973).
[20] P. Varilly, A.P. Willard, J.B. Kirkegaard, T.P.J. Knowles and D. Chandler, J. Chem. Phys. **146**, 135102 (2017).
[21] F. Müller-Plathe, ChemPhysChem **3**, 754–769 (2002).
[22] D. Reith, M. Putz and F. Müller-Plathe, J. Comput. Chem. **24**, 1624–1636 (2003).
[23] P. Carbone, H.A.K. Varzaneh, X.Y. Chen and F. Müller-Plathe, J. Chem. Phys. **128**, 064904 (2008).
[24] W.L. Jorgensen, J.D. Madura and C.J. Swenson, J. Am. Chem. Soc. **106**, 6638–6646 (1984).
[25] W.L. Jorgensen, D.S. Maxwell and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225–11236 (1996).
[26] Tinker – Software Tools for Molecular Design . <http://dasher.wustl.edu/tinker/>.
[27] G. Moad, Prog. Polym. Sci. **24**, 81–142 (1999).
[28] G. Moad, Prog. Polym. Sci. **24**, 1527–1528 (1999).
[29] Packmol . <http://m3g.iqm.unicamp.br/packmol/home.shtml>.
[30] L. Martínez, R. Andrade, E.G. Birgin and J.M. Martínez, J. Comp. Chem. **30**, 2157–2164 (2009).
[31] LAMMPS Molecular Dynamics Simulator . <http://lammps.sandia.gov>.
[32] S. Plimpton, J. Comp. Phys. **117**, 1–19 (1995).
[33] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic Press, San Diego, 2001).
[34] M.P. Allen and D.J. Tildesley, *Computer simulation of liquids*, 2nd ed. (Oxford University Press, Oxford, 2016).
[35] G.S. Grest and K. Kremer, Phys. Rev. A **33**, 3628–3631 (1986).
[36] G.S. Grest, K. Kremer and T.A. Witten, Macromolecules **20**, 1376–1383 (1987).
[37] G.S. Grest, Macromolecules **27**, 3493–3500 (1994).
[38] S.W. Sides, G.S. Grest and M.J. Stevens, Macromolecules **35**, 566–573 (2002).
[39] J.S. Pedersen and P. Schurtenberger, Macromolecules **29**, 7602–7612 (1996).
[40] W. Burchard, Adv. Polym. Sci. **48**, 1–124 (1983).
[41] B. Hammouda, Adv. Polym. Sci. **106**, 87–133 (1993).
[42] J.S. Higgins and H.C. Benoit, *Polymers and Neutron Scattering* (Clarendon Press, Oxford, 1997).
[43] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Clarendon Press, Oxford, 2002).
[44] S.M. Bhattacharjee, A. Giacometti and A. Maritan, J. Phys.: Condens. Matter **25**, 503101 (2013).
[45] M. Fixman, J. Chem. Phys. **36**, 306–310 (1962).
[46] O. Kratky and G. Porod, Recl. Trav. Chim. Pay B **68**, 1106–1122 (1949).